

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/111659/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Krumhuber, Eva G, Lai, Yukun ORCID: <https://orcid.org/0000-0002-2094-5680>, Rosin, Paul ORCID: <https://orcid.org/0000-0002-4965-3884> and Hugenberg, Kurt 2019. When facial expressions do and do not signal minds: the role of face inversion, expression dynamism, and emotion type. *Emotion* 19 (4) , pp. 746-750. 10.1037/emo0000475 file

Publishers page: <http://dx.doi.org/10.1037/emo0000475>  
<<http://dx.doi.org/10.1037/emo0000475>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



When facial expressions do and do not signal minds:  
The role of face inversion, expression dynamism, and emotion type

Eva G. Krumhuber

University College London

Yu-Kun Lai

Cardiff University

Paul L. Rosin

Cardiff University

&

Kurt Hugenberg

Miami University

Corresponding author: Eva Krumhuber, Department of Experimental Psychology,  
University College London, 26 Bedford Way, London WC1H 0AP, United Kingdom. E-mail:  
e.krumhuber@ucl.ac.uk

Authors' Note: We thank Matilda Hultgren, Benedek Paskuj, Toby Pilditch, Jingyan Li,  
Bishata Kandage, and Rhea Badwal.

Word Count: 2500

**© 2018, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors permission. The final article will be available, upon publication, via its DOI: 10.1037/emo0000475**

### **Abstract**

Recent research has linked facial expressions to mind perception. Specifically, Bowling and Banissy (2017) found that ambiguous doll-human morphs were judged as more likely to have a mind when smiling. Herein, we investigate three key potential boundary conditions of this “expression-to-mind” effect. First, we demonstrate that face inversion impairs the ability of happy expressions to signal mindful states in static faces; however, inversion does not disrupt this effect for dynamic displays of emotion. Finally, we demonstrate that not all emotions have equivalent effects. Whereas happy faces generate more mind ascription compared to neutral faces, we find that expressions of disgust actually generate less mind ascription than those of happiness.

*Keywords:* emotion; facial expression; configural processing; mind; morality; animacy

## **When facial expressions do and do not signal minds:**

### **The role of face inversion, expression dynamism, and emotion type**

Facial expressions of emotion can serve as “windows to the mind,” signaling others’ inner experiences, desires, and intents (Fridlund, 1994). Bowling and Banissy (2017) recently extended this work to demonstrate that emotional expressions can also signal whether targets have minds at all. Indeed, they demonstrated an “expression-to-mind link,” showing that smiling faces were more likely to be judged as having minds, relative to non-expressive (neutral) faces. Herein, we seek to replicate and extend Bowling and Banissy’s research by testing key boundary conditions that may influence how and when emotional expressions influence judgments of targets’ minds.

First, past research has reliably demonstrated that configural face information influences both the identification of facial expressions (e.g., Bombari et al., 2013; Prkachin, 2003) and mind perception (Deska, Almaraz, & Hugenberg, 2016; Hugenberg et al., 2016). Disrupting facial configurations by inverting faces both reduces emotion recognition abilities and leads to targets being rated as having less sophisticated minds. Given the joint role of face orientation in emotion perception and mind attribution, of primary interest was whether the previously established expression-to-mind link occurred even when faces were inverted, thereby hindering configural face processing. Thus, would emotional expressions have stronger effects on mind ascription when seen in upright versus inverted faces?

Second, *dynamic* (i.e., moving) facial expressions facilitate emotion recognition, in part because humans are sensitive to the direction in which expressions unfold (Krumhuber, Kappas, & Manstead, 2013). Similarly, moving faces may provide stronger cues to animacy than static faces. If inversion disrupts the expression-to-mind link for static faces, would this also occur for dynamic expressions? Given the richer expressive signal from dynamic faces,

we hypothesize that effects of inversion will be attenuated for dynamic displays of facial emotions.

Finally, whereas Bowling and Banissy (2017) showed that emotionally expressive targets are more likely to be ascribed sophisticated minds, this past work only tested happy versus neutral expressions. Of additional interest is whether other emotional expressions, and especially *negative* emotional expressions, signal that targets possess minds. Past research has linked mind perception to observers' desire for social affiliation (Powers, Worsham, Freeman, Wheatley, & Heatherton, 2014). Thus, perhaps the expression-to-animacy link is clear when faces signal affiliation (i.e., happy expressions). But would similar effects occur for disgust expressions? Because disgust signals revulsion (Rozin, Haidt, & McCauley, 2008), perhaps disgust expressions may be seen as *less* mentally sophisticated than neutrally-expressive faces.

### **Current Research**

In the current research, we sought to replicate and extend Bowling and Banissy's expression-to-mind link by testing 1) whether this link relies on configural processing, 2) whether dynamic signals of expressions override disruptions of configural processing, and 3) whether the expression-to-mind link occurs for both accepting expressions (happiness) and rejecting expressions (disgust).

To investigate these questions, participants judged whether human/cartoon morphs had minds, and we manipulated a) the orientation of faces (upright vs. inverted), b) the type of display (static vs. dynamic), and c) the emotional expression (happy vs. neutral vs. disgusted). We operationalized mind perception in two ways. Our primary measure was the ascription of *agentic mind* – whether targets can “think and plan” (Gray, Gray, & Wegner, 2007). Given that agentic minds also qualify entities as moral agents (i.e., capable of doing

good and evil; Gray, Young, & Waytz, 2012) a secondary measure of *moral agency* assessed whether targets seem “capable of moral action.”

## Method

**Participants and Design.** 283 students (215 women,  $M_{\text{age}}=22.2$ ,  $SD=5.75$ ) were recruited, ensuring 98% power to detect a 2 x 2 x 3 interaction effect (Cohen’s  $f = .10$ ,  $\alpha = .05$  two-tailed), assuming a 0.8 correlation between the measures. Because the target faces were White, only White/Caucasian participants were recruited to eliminate cross-race effects (Krumhuber, Swiderska, Tskankova, Kamble, & Kappas, 2015). The design included within-subject manipulations of emotion (3: disgust, happiness, neutrality) and morph level (7), and between-subject manipulations of display (2: static, dynamic) and orientation (2: upright, inverted). Participants were randomly assigned to one of the four between-subjects conditions, resulting in ~70 participants per condition. Participants received either course credit or £2 for participation. The study received ethical approval from the Department of Experimental Psychology, University College London, United Kingdom.

**Materials.** Stimuli consisted of 3 White male face identities (front view), each expressing the three emotions: disgust, happiness, and neutral. All expressions were obtained from the Amsterdam Dynamic Facial Expression Set (ADFES; Van der Schalk, Hawk, Fischer, & Doosje, 2011; see also Krumhuber, Skora, Küster, & Fou, 2017). We employed both static and dynamic portrayals of each expression. Dynamic stimuli were comprised of short video-clips (6s) which showed the face changing from non-expressive to peak emotional display. Static stimuli consisted of a single frame of the peak expression. For neutral displays, no emotional expression was visible, although the videos showed minor naturalistic movements of targets’ head and eyes.

To create the faces’ artificial analogues, modifications were applied to the facial texture in each image or video frame (25fps; see Rosin & Lai, 2015). This produced realistic

cartoons of the same identities. We next used the human and cartoon versions of each image to create human/cartoon morphs; human and cartoon facial stimuli were morphed parametrically in seven equidistant steps of 16.7% (see Figure 1), netting 63 static and 63 dynamic stimuli (3 faces x 3 emotions x 7 morphs). We further manipulated face orientation (upright vs. inverted) for each display, resulting in 252 stimuli. All stimuli were displayed in color on white backgrounds.

**Procedure and Measures.** After providing informed consent, participants were informed that they would see faces that varied in appearance from human to artificial. Their task was to rate the extent to which they believed each target “has a mind” and “can act morally.” Ratings of mind and moral agency were blocked; with order counterbalanced across participants.

In each of the two blocks, participants were presented with the same set of 21 randomly sequenced face stimuli at the center of the screen (6s per face). Face stimuli showed the three emotions (disgust, happy, neutral), each expressed by a different face identity, at seven morph levels (the same target only portrayed one emotion in a set). The face identity-emotion mapping was counterbalanced between-subjects. After viewing each stimulus, participants provided ratings for perceived mind and moral agency using a 7-point Likert scale (1=*definitely does not have a mind/cannot act morally*; 7=*definitely has a mind/can act morally*) before viewing and rating the next stimulus.

## Results

Ratings for agentic mind and moral agency were linearly transformed from the original 7-point scale into scores ranging from 0 to 1 (with higher scores indicating greater levels of the dimension). They were then averaged across the three face identities and submitted to 3 (emotion: disgust, happiness, neutrality) x 7 (morph level) x 2 (display: static, dynamic) x 2 (orientation: upright, inverted) ANOVAs, with the latter two variables between-

subjects. The ANOVAs yielded a three-way interaction between display, orientation, and emotion for both ratings of mind,  $F(2, 558) = 3.08, p = .048, \eta_p^2 = .01$ , observed power = .59; and moral agency,  $F(2, 558) = 4.66, p = .011, \eta_p^2 = .02$ , observed power = .76.<sup>1</sup> To decompose the interactions, separate ANOVAs with emotion, morph level, and orientation were conducted for static and dynamic displays. For all analyses, a Greenhouse-Geisser adjustment to degrees of freedom was applied and Bonferroni correction was used for multiple comparisons (see Supplementary Materials for model fitting and PSE analyses).

### Static displays

The analysis revealed main effects of morph level for ratings of mind,  $F(6, 846) = 414.32, p < .001, \eta_p^2 = .75$ , and moral agency,  $F(6, 846) = 311.64, p < .001, \eta_p^2 = .69$ . Facial stimuli were ascribed higher levels of mind and moral agency the more human-like they looked. There were also main effects of emotion for mind perception,  $F(2, 282) = 5.14, p = .007, \eta_p^2 = .03$ , and moral agency,  $F(2, 282) = 31.88, p < .001, \eta_p^2 = .18$ . However, these were qualified by a significant interaction between emotion and orientation: mind,  $F(2, 282) = 5.02, p = .008, \eta_p^2 = .03$ , observed power = .80; moral agency,  $F(2, 282) = 8.34, p < .001, \eta_p^2 = .06$ , observed power = .95.

Emotional expressions exerted a greater impact in the perception of upright than inverted faces, see Fig 2a. Upright faces received lower ratings of mind when expressing disgust relative to either neutrality,  $t(72) = -3.22, p = .006, 95\% \text{ CI } [-0.09, -0.01], d = -.29$ , or happiness,  $t(72) = -4.78, p < .001, 95\% \text{ CI } [-0.09, -0.03], d = -.36$ . No such difference between emotions was observed for inverted faces,  $ps > .99, ds < .01$ .

The pattern of results was similar for ratings of moral agency. Disgust led in upright faces to lower ascriptions of moral ability than did neutrality,  $t(72) = -6.26, p < .001, 95\% \text{ CI } [-0.14, -0.06], d = -.59$ , or happiness,  $t(72) = -6.70, p < .001, 95\% \text{ CI } [-0.18, -0.09], d = -.79$ . Happy expressions also differed significantly from neutral expressions,  $t(72) = 2.45, p = .032$ ,



95% CI [0.002, 0.08],  $d = .24$ . For inverted faces, a distinction was only made between disgust and neutral expressions,  $t(69) = -2.75$ ,  $p = .029$ , 95% CI [-0.08, -0.003],  $d = -.25$ . All other comparisons were non-significant,  $ps > .09$ ,  $ds < .02$ .<sup>2</sup>

### Dynamic displays

The analyses revealed main effects of morph level for both ratings of mind,  $F(6, 828) = 282.28$ ,  $p < .001$ ,  $\eta_p^2 = .67$ , and moral agency,  $F(6, 828) = 207.68$ ,  $p < .001$ ,  $\eta_p^2 = .60$ , such that faces were ascribed greater mental and moral capacity with increasingly human-like appearance. In addition, there were main effects of emotion for mind perception,  $F(2, 276) = 22.81$ ,  $p < .001$ ,  $\eta_p^2 = .14$ , and moral agency,  $F(2, 276) = 47.14$ ,  $p < .001$ ,  $\eta_p^2 = .25$ . As seen in Figure 2b, mind attributions were lower in response to neutral expressions relative to either disgust,  $t(139) = 3.01$ ,  $p = .009$ , 95% CI [0.01, 0.06],  $d = .19$ , or happiness,  $t(139) = 6.54$ ,  $p < .001$ , 95% CI [0.04, 0.10],  $d = .42$ . Also, happy expressions significantly differed from those of disgust,  $t(139) = -3.84$ ,  $p = .001$ , 95% CI [-0.06, -0.01],  $d = -.23$ .

For ratings of moral agency, disgust led to lower attributions compared to either neutrality,  $t(139) = -3.32$ ,  $p = .004$ , 95% CI [-0.08, -0.01],  $d = -.31$ , or happiness,  $t(139) = -8.88$ ,  $p < .001$ , 95% CI [-0.17, -0.10],  $d = -.84$ . Happiness also received higher ratings than did neutrality,  $t(139) = 7.03$ ,  $p < .001$ , 95% CI [0.06, 0.11],  $d = .55$ .

The interaction between emotion and orientation was non-significant: mind,  $F(2, 276) = 0.64$ ,  $p = .529$ ,  $\eta_p^2 = .005$ , observed power = .16; morality,  $F(2, 276) = 1.92$ ,  $p = .152$ ,  $\eta_p^2 = .014$ , observed power = .38. Thus, face orientation did not modulate the effects of emotion in dynamic faces.

### Discussion

Herein, we sought to replicate and extend the previously demonstrated expression-to-mind link. We conceptually replicated Bowling and Banissy's (2017) previous findings that, in most of our conditions, happy faces were seen as having more sophisticated minds than

neutrally-expressive faces. However, we also found that face inversion attenuated the effects of emotional expressions on mind perception in static faces. Consistent with previous evidence (e.g. Prkachin, 2003), configural face processing seems necessary for static displays of emotion to affect mind perception. Notably, however, dynamic faces showed different results. Face inversion did *not* hinder the effects of expressions on mind perception. Thus, Bowling and Banissy's original finding of happy faces appearing more mindful was equally strong for upright and inverted dynamic displays.

Second, the current data demonstrate that not all emotions equivalently signal mind. Whereas upright happy expressions were typically seen as having more agentic minds and greater moral agency, disgusted expressions tended to show the opposite pattern. Consistent with past research (e.g., Powers et al., 2014), perceivers are more likely to attribute minds to targets signaling social acceptance (smiling) than social rejection (disgust). Notably, this disgust-driven reduction of mind perception is observable for both measures of agentic minds. Further, an additional study (see Supplementary Materials) replicates the findings using both new measures and naturalistic expressions of emotions, further supporting these differential effects of expressions on mind attribution. Somewhat surprising was that dynamic neutral faces attracted lowest ratings of mind. This could be due to the perceived lack of behavioral intentions as evident in a moving, but non-expressive face. Hence, it is possible that valence may play a stronger role in judging whether a target is capable of moral agency (i.e., good vs. bad actions) compared to whether targets have minds (i.e., are mentally capable vs. incapable). Future research would benefit from investigating this in more detail.

Taken together, the present research conceptually replicates previous findings showing that happy faces trigger more mind ascription. Here, we also show that these effects are bounded by face orientation, expression dynamism, and emotion type.

## References

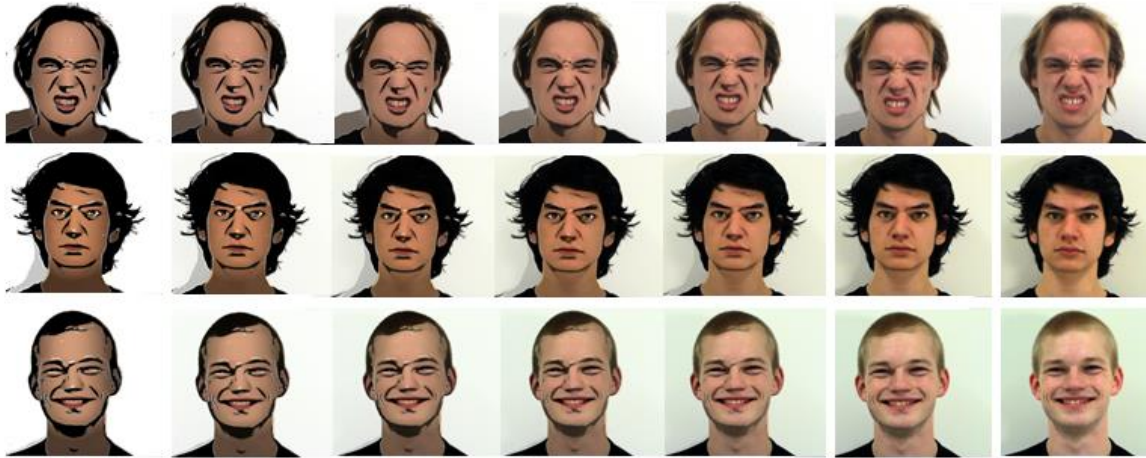
- Bombari, D., Schmid, P. C., Mast, M. S., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. *The Quarterly Journal of Experimental Psychology*, 66, 2426-2442.
- Bowling, N. C., & Banissy, M. J. (2017). Emotion expression modulates perception of animacy from faces. *Journal of Experimental Social Psychology*, 71, 83-95.
- Deska, J. D., Almaraz, S. M., & Hugenberg, K. (2016). Of mannequins and men: Ascriptions of mind from faces are bounded by perceptual and processing similarities to human faces. *Social Psychological and Personality Science*, 42, 1666-1677.
- Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. San Diego: Academic Press.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315, 619.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry*, 23, 101-124.
- Hugenberg, K., Young, S. G., Rydell, R. J., Almaraz, S. M., Stanko, K. A., See, P. E., & Wilson, J. P. (2016). The face of humanity: Configural face processing influences ascriptions of humanness. *Social Psychological and Personality Science*, 7, 167-175.
- Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you are doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, 90, 543-555.
- Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review*, 5, 41-46.
- Krumhuber, E. G., Skora, L., Küster, D., & Fou, L. (2017). A review of dynamic datasets for facial expression research. *Emotion Review*, 9, 280-292.

- Krumhuber, E. G., Swiderska, A., Tskankova, E., Kamble, S. V., & Kappas, A. (2015). Real or artificial? Intergroup biases in mind perception in a cross-cultural perspective. *PlosOne*, 10(9), e0137840.
- Looser, C. E., & Wheatley, T. (2010). The tipping point of animacy. *Psychological Science*, 21, 1854-1862.
- Powers, K. E., Worsham, A. L., Freeman, J. B., Wheatley, T., & Heatherton, T. F. (2014). Social connection modulates perceptions of animacy. *Psychological Science*, 25, 1943-1948.
- Prkachin, G. C. (2013). The effects of orientation on detection and identification of facial expressions of emotion. *British Journal of Psychology*, 94, 45-62.
- Rosin, P. L., & Lai, Y. K. (2015). Non-photorealistic rendering of portraits. *Proceedings of the Workshop on Computational Aesthetics* (pp. 159-170). New York, USA: ACM Press.
- Rozin, P., Haidt, J., & McCauley, C. R. (2008). Disgust. In M. Lewis, J. M., Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions*, 3<sup>rd</sup> ed. (pp. 757-776). New York: Guilford Press.
- Van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving faces, looking places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion*, 11, 907-920.
- Wallhoff, F. (2004). FGnet – Facial expression and emotion database. [Online]. Retrieved from <http://www-prima.inrialpes.fr/FGnet/html/benchmarks.html>

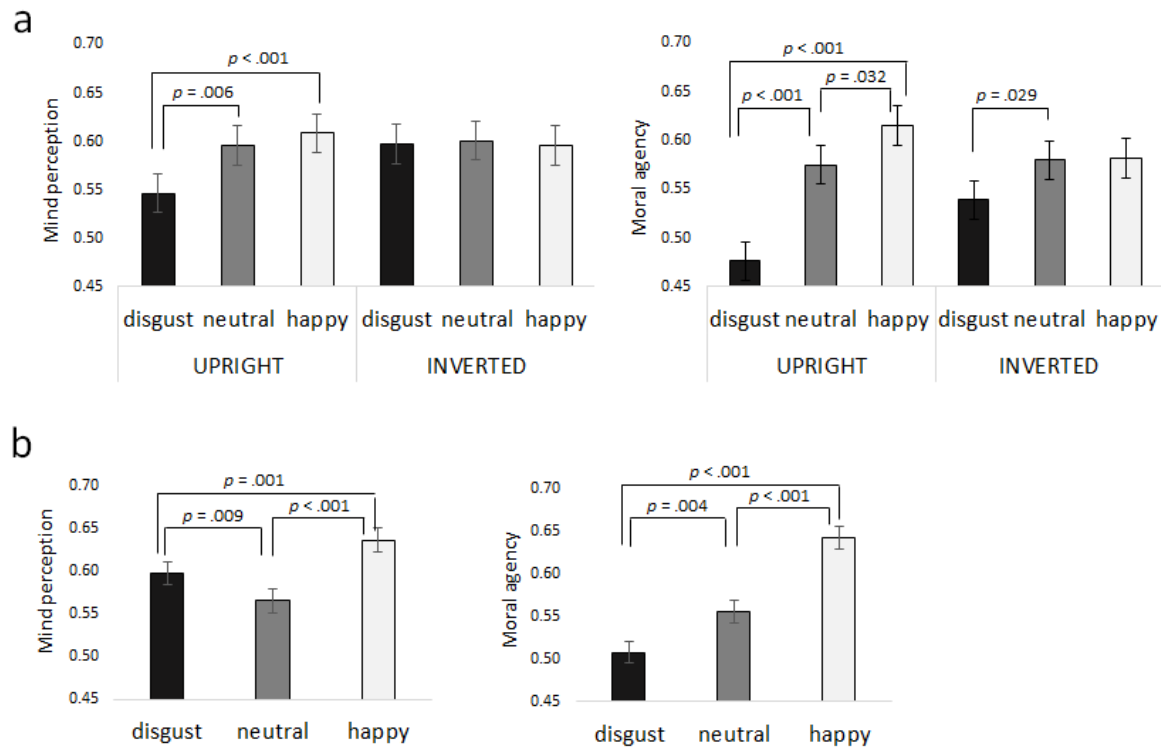
### Footnotes

<sup>1</sup> There were also significant main effects of morph level,  $F_{\text{mind}}(6, 1674) = 684.20, p < .001, \eta_p^2 = .71$ ,  $F_{\text{morality}}(6, 1674) = 513.36, p < .001, \eta_p^2 = .65$ ; and emotion,  $F_{\text{mind}}(2, 558) = 16.10, p < .001, \eta_p^2 = .06$ ,  $F_{\text{morality}}(2, 558) = 75.24, p < .001, \eta_p^2 = .21$ ; morph level by display interactions,  $F_{\text{mind}}(6, 1674) = 3.26, p = .035, \eta_p^2 = .01$ ,  $F_{\text{morality}}(6, 1674) = 5.12, p = .005, \eta_p^2 = .02$ ; emotion by display interactions,  $F_{\text{mind}}(2, 558) = 12.36, p < .001, \eta_p^2 = .04$ ,  $F_{\text{morality}}(2, 558) = 6.51, p = .002, \eta_p^2 = .02$ ; emotion by morph level interactions,  $F_{\text{mind}}(12, 3348) = 1.90, p = .035, \eta_p^2 = .01$ ,  $F_{\text{morality}}(12, 3348) = 2.21, p = .011, \eta_p^2 = .01$ ; and emotion by orientation interactions,  $F_{\text{morality}}(2, 558) = 4.59, p = .012, \eta_p^2 = .02$ . Because the main effects and lower-order interactions were qualified by a significant three-way interaction between display, orientation, and emotion, we focused on the highest order interaction.

<sup>2</sup> For ratings of moral agency, there was also a significant interaction between emotion and morph level:  $F(12, 1692) = 2.72, p = .002, \eta_p^2 = .02$ . For all seven morph levels, disgust expressions were ascribed less moral agency relative to happy ( $ps < .01$ ) and neutral expressions ( $ps < .05$ ). In addition, happiness led to higher ratings of moral agency compared to neutrality in morph 4 ( $p = .011$ ).



*Figure 1.* Facial exemplars of three male targets ranging from artificial (left) to human-realistic (right), showing disgust (top row), no emotion/neutral (middle row), and happiness (bottom row).



*Figure 2.* Ratings of mind and moral agency for disgust, neutral, and happy expressions as a function of (a) orientation and emotion in static displays, and (b) emotion in dynamic displays. Error bars represent SEM.

## **Supplementary Materials**

When facial expressions do and do not signal minds:  
The role of face inversion, expression dynamism, and emotion type

by E. G. Krumhuber, Y. Lai., P. L. Rosin, & K. Hugenberg

### **Model fitting**

Average ratings of mind and moral agency were obtained across the three facial exemplars for each point on the morph continuum (from 1- artificial to 7- human). The resulting mean values were then linearly transformed from the original 7-point Likert scale to scores between 0 and 1 (with higher scores indicating greater levels of perceived mind and moral agency). To obtain psychometric curves for each dependent measure, the standardized scores were fitted with a Gaussian distribution in GraphPad Prism 6 (GraphPad Software Inc., California, USA). This provided an overall fit index of participants' judgment data to the mean estimated slope. As shown in Figures S1 and S2, a good fit was achieved for all models ( $r^2s \geq .86$ ) which allowed for the calculation of the Point of Subjective Equality (PSE). Outliers were identified on the individual PSE values for each participant falling beyond the  $M \pm 2.5 SD$  range and were treated as missing data in the PSE analyses.



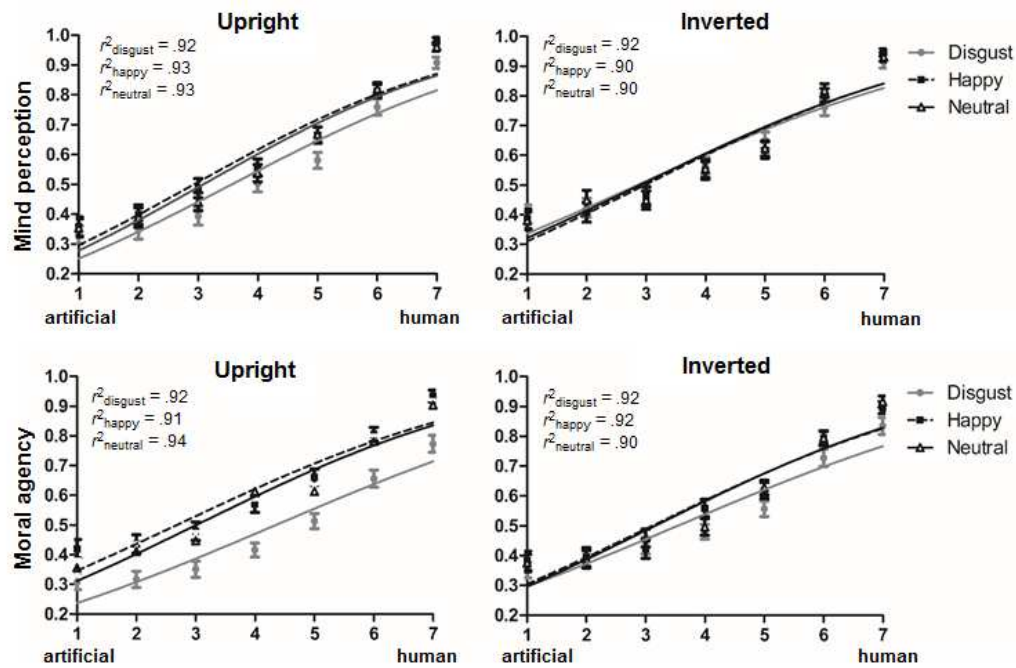
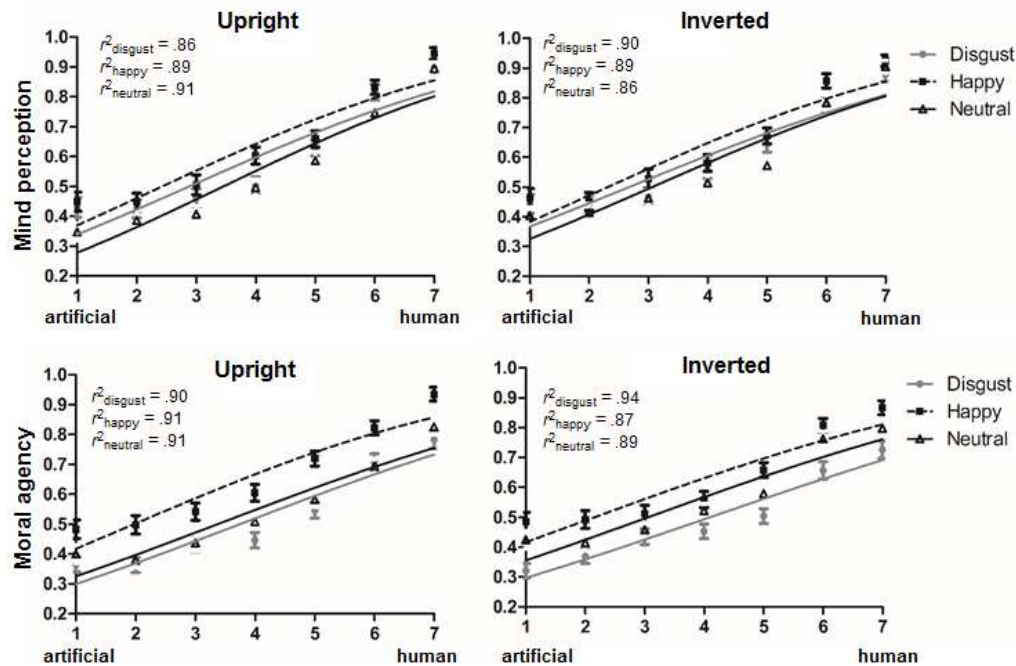


Figure S1. Mean ratings of mind and moral agency for disgust, happy, and neutral expressions in static displays at each point along the morph continuum (1-7) including error bars (SEM) and the fitted curves per measure.  $r^2$  = model fit index.



*Figure S2.* Mean ratings of mind and moral agency for disgust, happy, and neutral expressions in dynamic displays at each point along the morph continuum (1-7) including error bars (SEM) and the fitted curves per measure.  $r^2$  = model fit index.

### **Point of Subjective Equality (PSE) analyses**

PSE values were derived from the fitted curves in GraphPad Prism 6 where the face ratings correspond to the ordinate midpoint of a measure (0.5 at y-axis in Figures S1 and S2), thereby reflecting the point on the morph continuum where the face is judged 50% animate. ANOVAs with the within-subjects factor emotion (disgust, happy, neutrality) and the between-subjects factor orientation (upright, inverted) were performed on the PSE scores of mind and moral agency separately for static and dynamic displays.

### **Static displays**

The analysis revealed a significant main effect of emotion for ratings of moral agency,  $F(2, 218) = 10.15, p < .001, \eta_p^2 = .09$ , but not mind,  $F(2, 234) = 0.67, p = .502, \eta_p^2 = .01$ . The emotion by orientation interaction was significant for both mind,  $F(2, 234) = 4.84, p = .011, \eta_p^2 = .04$ , and moral agency,  $F(2, 218) = 4.78, p = .016, \eta_p^2 = .04$ . As shown in Figure S3a, thresholds for attributing mind,  $t(66) = 1.98, p = .023, 95\% \text{ CI } [0.08, 1.40], d = .22$ , and moral agency,  $t(59) = 4.01, p < .001, 95\% \text{ CI } [1.04, 3.67], d = .72$ , were significantly higher for upright faces displaying disgust compared to happiness. That is, disgust faces needed to look more human-like in order for them to be seen as mindful and morally agentic than happy faces. Disgust expressions also led in upright faces to elevated thresholds in moral agency ascription compared to neutral expressions,  $t(58) = 3.51, p = .001, 95\% \text{ CI } [0.43, 2.25], d = .49$ , which in turn were judged at higher morph levels than happiness,  $t(61) = 1.81, p$

= .019, 95% CI [-1.89, -0.13],  $d = .29$ . No such difference between emotions in mind and morality thresholds was observed for inverted faces,  $ps > .52$ ,  $ds < .17$ .

### Dynamic displays

The analysis revealed a significant main effect of emotion for ratings of mind,  $F(2, 222) = 11.30$ ,  $p < .001$ ,  $\eta_p^2 = .09$ , and moral agency,  $F(2, 208) = 10.30$ ,  $p < .001$ ,  $\eta_p^2 = .09$ . As seen in Figure S3b, thresholds for attributing mind were higher in response to neutral expressions relative to either disgust,  $t(117) = -3.16$ ,  $p = .021$ , 95% CI [-1.16, -0.07],  $d = -.26$ , or happiness,  $t(119) = -5.13$ ,  $p < .001$ , 95% CI [-1.41, -0.46],  $d = -.44$ . Happiness led to the lowest thresholds in morality attributions compared to neutrality,  $t(110) = -3.42$ ,  $p = .003$ , 95% CI [-2.02, -0.33],  $d = -.35$ , and disgust,  $t(112) = 4.57$ ,  $p < .001$ , 95% CI [0.70, 2.71],  $d = .58$ . The interaction between emotion and orientation was non-significant: mind,  $F(2, 222) = 0.05$ ,  $p = .946$ ,  $\eta_p^2 = .00$ ; morality,  $F(2, 208) = 0.94$ ,  $p = .390$ ,  $\eta_p^2 = .01$ . Thus, face orientation did not modulate the effects of emotion on perception thresholds in dynamic faces.

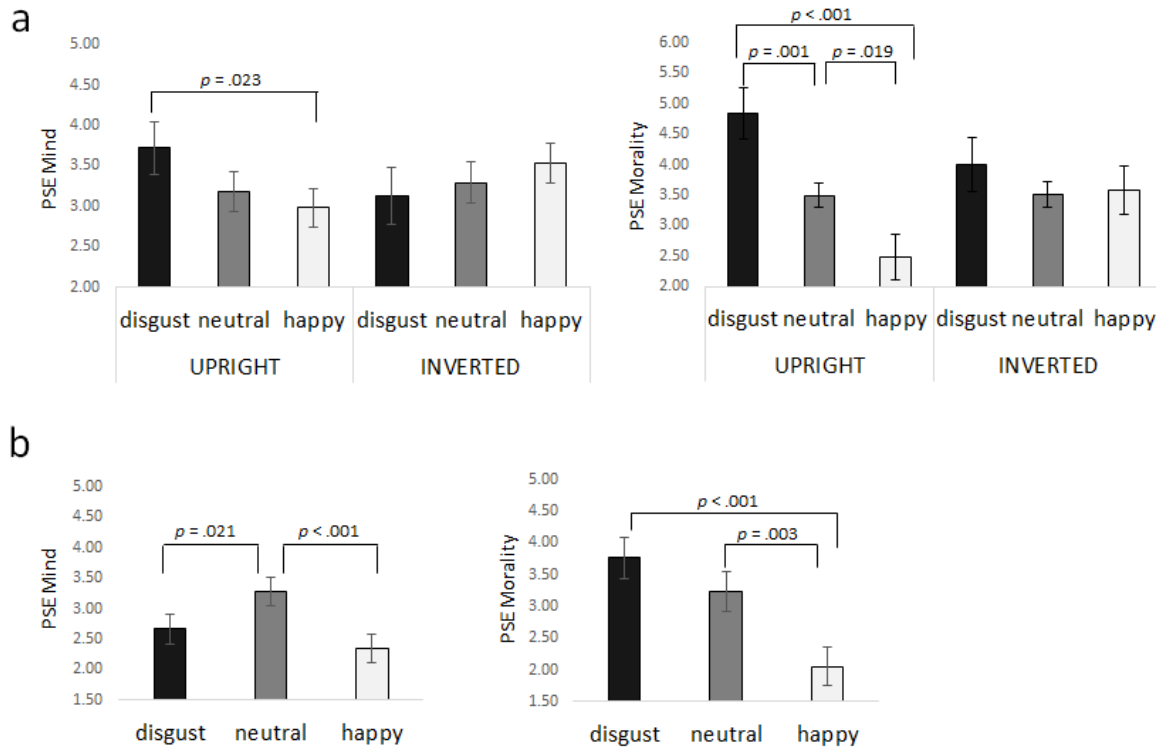


Figure S3. Mean PSE values (shown as original morph numbers (1-7)) of mind and moral agency for disgust, neutral, and happy expressions as a function of (a) orientation and emotion in static displays, and (b) emotion in dynamic displays. Error bars represent SEM.

### Additional Replication of the Differential Effects of Expressions on Mind Ascription

In order to replicate the effect of emotion on mind perception with more naturalistic stimuli and a new measure of mind ascription, 77 participants (36 women,  $M_{age} = 38.4$  years,  $SD = 13.1$ ) were recruited via Amazon's Mechanical Turk. These participants saw a series of still images of naturalistic expressions, varying in emotion (disgusted, happy, neutral). Images were taken from the FGnet database (Wallhoff, 2004; for a review see Krumhuber et al., 2017) and presented in color (size: 640 x 480 pixels). Stimuli depicted three White male identities, each displaying a disgust, happy, or neutral expression under spontaneous emotion-eliciting conditions. Participants saw only one exemplar per emotion, thereby

counterbalancing the face identity-emotion mapping between participants. For each stimulus, participants completed the 10-item Mind Attribution Scale (Kozak, Marsh, & Wegner, 2006,  $\alpha = 0.92$ ) which was presented on the same screen. Scale ratings for each target were averaged across expression type and subjected to a repeated-measures ANOVA.

This ANOVA revealed a significant main effect of emotion on mind perception,  $F(2, 152) = 3.26, p = .041, \eta_p^2 = .04$ . Overall, disgust resulted in lower ascriptions of mind compared to happiness,  $t(76) = 2.61, p = .033, 95\% \text{ CI } [0.01, 0.44], d = .22$ . Although the difference between disgust vs. neutral and happy vs. neutral was not significant ( $ps > .23, ds < .19$ ), a significant linear trend emerged for emotion,  $F(1, 76) = 6.81, p = .011, \eta_p^2 = .08$ , suggesting that mind attributions increased with the positive valence of the expression (disgust < neutral < happiness).